Post-Marketing Surveillance for Adverse Drug Reaction using Clinical Data

Sujay Bankar,Suyog Gaikwad,Sidhesh Gawas,Mayur Shirkare,Mrs.Deepa Abin

***Abstract*— Adverse drug reactions (ADRs) is a big challenge in drug development process. Medicines are designed to cure, treat, or prevent diseases; however, there are also risks in taking any medicine. Particularly short term or long term ADR's can cause serious harm to patients. ADR's are the harmful reactions of the drugs caused to humans due to allergies, overdose, chemical reactions between two chemicals in the medicines, etc. Discovering unknown ADR's in post-marketing surveillance as early as possible is of great importance, as it save lives and prevent harmful consequences. We will be using data mining technique Naive Bayes classifier for classification of ADR's. Timely safety surveillance after a drugs release on the market is therefore an urgent goal of public health systems. We are going to propose software system approach for proactive monitoring and detecting potential ADR's using clinical records.**

Index Terms*— *Adverse Drug Reaction (ADR), Post-marketing surveillance, Naïve Bayes, Association Rule Mining (ARM)

I. INTRODUCTION

ADRs represent a major world-wide problem [1]. They can complicate a patient's medical condition and contribute to increased unhealthiness, even death. The premarketing clinical trials are required for all drugs available in the market, which are limited in sample-size and duration, and hence they are not capable of rare ADRs detection. Drug safety is dependent heavily on post marketing surveillance, i.e. the practice of Monitoring pharmaceutical drugs.

An adverse drug reaction is a "response to a drug which is Noxious and unintended. It occurs at doses normally Used in humans for prophylaxis, diagnosis, or therapy of disease or for the modification of physiologic function. It is very important to understand that there is a link between a drug and an adverse drug reaction. An allergy is an adverse drug reaction mediated by an immune response (e.g., rash, hives). A side effect is an expected and known effect of a drug that is not the intended therapeutic outcome. The term "side effect" tends to normalize the concept of injury from drugs. There has been recommendation that this term should generally be avoided in favor of adverse drug reaction.

The idea behind the work is to minimize the occurrence of ADR's by processing the past medical data by using Naïve Bayes and Association Rule Mining.

II. PROCEDURE FOR PAPER SUBMISSION

A. Review Stage

Submit your manuscript electronically for review.

B. Final Stage

When you submit your final version, after your paper has been accepted, prepare it in two-column format, including figures and tables.

C. Figures

As said, to insert images in *Word*, position the cursor at the insertion point and either use Insert | Picture | From File or copy the image to the Windows clipboard and then Edit | Paste Special | Picture (with "Float over text" unchecked).

The authors of the accepted manuscripts will be given a copyright form and the form should accompany your final submission.

III. LITERATURE SURVEY

The traditional measures used for detecting the ADRs are complicated and have many risks. The drawback of these measures were that they were used for premarketing surveillance which is not effective to detect rare ADRs [1]. To overcome this drawback, YanqingJi has proposed a new approach called as Post-marketing surveillance [3].The Post-Marketing Surveillance (PMS) is an effective method to detect ADR's. It follows spontaneous reporting, so actions

Manuscript received Jan, 2016.

***Sujay Bankar**, Computer, Pimpri Chinchwad College Of Engineering, ., Pune, INDIA, 9665988415*

***Suyog Gaikwad**, Computer, Pimpri Chinchwad College Of Engineering,, Pune, INDIA, 8446373511.*

***Sidhesh Gawas**, Computer, Pimpri Chinchwad College Of Engineering,, Pune, INDIA,9922703286.*

***Mayur Shirkare**, Computer, Pimpri Chinchwad College Of Engineering, Pune, INDIA, 7875253298.*

***Mrs.Deepa Abin**, Computer, Pimpri Chinchwad College Of Engineering, Pune, INDIA, 9226094976.*

ISSN: 2278 – 1323

taken to manage new ADR's are faster. To make the system more efficient we are focusing on the method proposed by Amiya Kumar Tripathi i.e. the combination of Naïve Bayes and Association Mining for fast classification of database and generating rule [1]

IV. OVERVIEW OF PROPOSED SYSTEM

The patient's data is an important aspect of the system which is a source of information required for processing of the data. Here, mining potential ADR signals is done through an extensive data source that contains administrative data, pharmacy, and clinical laboratory data.

The proposed model consists of the following important concepts:

a) Acquisition of data
b) Preprocessing
c) Naïve Bayes based classification
d) Detection and Prediction
e) Association Rule mining

a) Acquisition of data:

We are going to collect the data (i.e. medical history of patient, percentage compositions of elements of medicine, etc) from the likes of pharmacists, MR's, Doctors, etc. The collected data may be in the structured format or unstructured format.

b) Preprocessing:

In preprocessing step, semi structured or unstructured data will be converted into the structured format by organizing the data according to appropriate attributes.

c) Classification:

The Naïve Bayesian model is simple and easy to build. It has no complicated parameters which makes it useful for large datasets. We are using Naïve Bayes classification for classifying data. This classification was named after the Thomas Bayes (1702-1761) who proposed Bayes theorem. The Bayesian classification provides practical learning algorithms and prior knowledge and observed data can be combined [2].

d) Detection And Prediction

By using Naive Bayes, we process patient's medical data to detect whether he/she is having ADR or not. By using Association rule we predict the appropriate rules to prevent the ADR [4].

e) Association Rule mining :

Association Rule mining is also easy and simple to execute. It was initially used for market basket analysis which included relating the customer purchases [3]. In several data mining tasks, frequent pattern mining plays an important role in generating association rules .[6]

Rules of model:

- An association rule is an implication of the form:

  $X \rightarrow Y$, where $X, Y \subset I$, and $X \cap Y = \varnothing$ [6]

- An item set is a set of items.
- E.g., X = {crocin, sarradon, paracetamol} is an item set.
- A *k*-itemset is an itemset with *k* items.
- E.g., {crocin, sarradon, paracetamol} is a 3-itemset.

Fig 1 below shows the overview of the system.

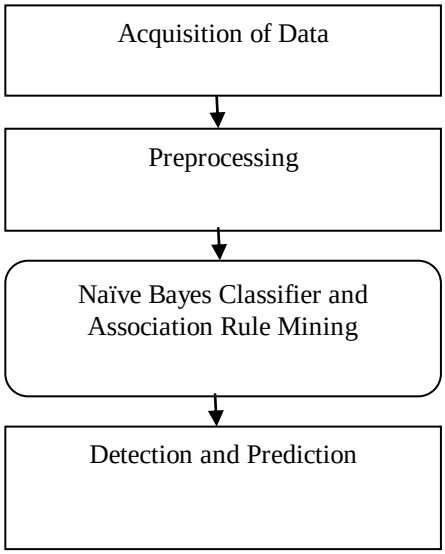

Fig 1 : System Model.

System will consist of four major components. They are Acquisition of data, Pre-processing, a, Naïve Bayes classifier and association rule mining, Detection and prediction. The component will be designed as functional unit and will be implemented

as classes. In acquisition of data the data sets are collected in unstructured format. Then in pre-processing step this unstructured data is processed and converted in structured XML format. By using Naive Bayes, we process patient's medical data to detect whether he/she is having ADR or not. By using Association rule we predict the appropriate rules to prevent the ADR.

V ALGORITHMS

Naïve Bayes:

Naïve Bayes (NB), a special form of Bayesian Network has been widely used for data classification in that its predictive performance is competitive with state-of-the-art classifiers [2]. As a classifier, it learns from training data from the conditional probability of each attribute given the class label. It uses Bayes rule to compute the probability of the classes given the particular instance of the attributes, prediction of the class is done by identifying the class with the highest posterior probability [5].Research shows naïve Bayes still performs well in spite of strong dependencies among attributes.

- Naïve Bayes Theory:-
- P(h/D)= P(D/h) P(h)
  P(D)
- P(h) : Prior probability of hypothesis h
- P(D) : Prior probability of training data D
- P(h/D) : Probability of h given D
- P(D/h) : Probability of D given h

Apriori Algorithm:

- Association rule generation is usually divided into two steps
- First, minimum support is applied to find all frequent item sets in a database [3].
- Second, these frequent item sets and the minimum confidence constraint are used to form rules [3].

VI PROPOSED MODEL

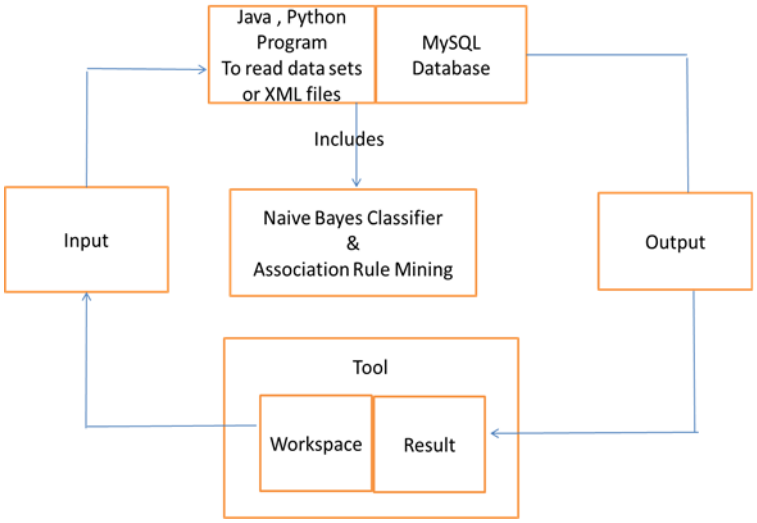

Fig2:Proposedmodel

VII.EDITORIAL POLICY

The submitting author is responsible for obtaining agreement of all coauthors and any consent required from sponsors before submitting a paper. It is the obligation of the authors to cite relevant prior work.

Authors of rejected papers may revise and resubmit them to the journal again.

VIII.PUBLICATION PRINCIPLES

The contents of the journal are peer-reviewed and archival. The journal INTERNATIONAL JOURNAL OF ADVANCED RESEARCH IN COMPUTE ENGINEERING & TECHNOLOGY (IJARCET) publishes scholarly articles of archival value as well as tutorial expositions and critical reviews of classical subjects and topics of current interest.

Authors should consider the following points:

1) Technical papers submitted for publication must advance the state of knowledge and must cite relevant prior work.
2) The length of a submitted paper should be commensurate with the importance, or appropriate to

ISSN: 2278 – 1323

the complexity, of the work. For example, an obvious extension of previously published work might not be appropriate for publication or might be adequately treated in just a few pages.

3) Authors must convince both peer reviewers and the editors of the scientific and technical merit of a paper; the standards of proof are higher when extraordinary or unexpected results are reported.
4) Because replication is required for scientific progress, papers submitted for publication must provide sufficient information to allow readers to perform similar experiments or calculations and use the reported results. Although not everything need be disclosed, a paper must contain new, useable, and fully described information. For example, a specimen's chemical composition need not be reported if the main purpose of a paper is to introduce a new measurement technique. Authors should expect to be challenged by reviewers if the results are not supported by adequate data and critical details.

IX. CONCLUSION

Adverse drug reactions have implications not only for the patient, but for the entire health care system. Reporting of ADRs provides clinicians and health care companies valuable insight into the toxicity profile of an agent. Many ADRs are preventable, although some effects cannot be avoided (e.g. nausea in chemo treatment for cancer) Better research and greater understanding of disease processes will lead to more effective prediction of ADR's. The proposed system will increase the rate of recognition of ADR's and thereby an attempt has been made to improve medical care.

ACKNOWLEDGMENT

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank ... ." Instead, write "F. A. Author thanks ... ." **Sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page**.

X. REFERENCES

[1] Detection of Adverse Drug Events through Data Mining Techniques by Amiya Kumar Tripathy, Nilakshi Joshi, Mario Durando, Loyal Carvalho, 2015.

[2] AESNB: Active Example Selection with Naïve Bayes Classifier for Learning from Imbalanced Biomedical Data by Min Su Lee, Byoung-Hee Kim, Byoung-Tak Zhang, 2009.

[3] A Multi-Relational Association Mining Algorithm for Screening Suspected Adverse Drug Reactions by Yanqing Ji, Fangyang Shen, John Tran, 2014.

[4] Data Mining to Generate Adverse Drug Events Detection Rules by Emmanuel Chazard, Gr´egoire Ficheur, 2011.

[5] Medical datamining with a new algorithm for Feature Selection and Naïve Bayesian classifier by Ranjit Abraham, Jay B.Simha, Iyengar S.S, 2007

[6]Mining Medical Databases using Proposed Incremental Association Rules Algorithm (PIA) by Laila Elfangary , Walid Adly Atteya.

[7]Causal Association Mining for Detection of Adverse Drug Reactions by Mrs. Deepa Abin, Tanushree C. Mahajan, Manali S. Bhoj, Swapnil Bagde,Dr. K. Rajeswari

Sujay Bankar pursuing Computer Enginerring degree in Pimpri Chinchwad College Of Engineering,
Suyog Gaikwad pursuing Computer Enginerring degree in Pimpri Chinchwad College Of Engineering,
Sidhesh Gawas pursuing Computer Enginerring degree in Pimpri Chinchwad College Of Engineering,
Mayur Shirkare pursuing Computer Enginerring degree in Pimpri Chinchwad College Of Engineering,
Mrs.Deepa Abin BE(Comp),M.E(Comp),Area of Interest:Data Mining